DOCUMENT RESUME

ED 082 384 EA 005 517

AUTHOR Levine, Daniel U.; And Others
TITLE An Exploration of the Use of Socioeconomic Census Data To Predict Achievement and Evaluate the Effects of Concentrated Urban Poverty Among Elementary Schools in a Big City. Final Report.
INSTITUTION Missouri Univ., Kansas City. Center for the Study of Metropolitan Problems in Education.
SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C. Regional Research Program.
BUREAU NO BR-1-G-012
PUB DATE 1 Feb 73
GRANT OEG-7-71-0006(509)
NOTE 43p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Academic Achievement; Achievement Tests; Census Figures; *Economic Disadvantagement; *Elementary Schools; *Predictor Variables; *Socioeconomic Influences; Statistical Analysis; Urban Areas

ABSTRACT

The major purpose of this study was to determine whether data from the 1970 Census could be used to account for at least 70 percent of the variance in grade-level achievement scores in the elementary schools of a big city, without including variables directly denoting race or ethnicity. This goal was attained when it was found that a regression equation, using four variables from the Third Count (block statistics) of the Census, accounted for 75 percent of the variance in 6th grade achievement scores in a sample of 122 elementary schools in Chicago. The four variables (Per Cent of Females Separated; Per Cent of Families Which Lack One or More Plumbing Facilities; Per Cent of Units With Six Persons or More; and Per Cent Owner Occupied Units) also provided the basis for exploratory analysis with a number of additional variables. Separate analysis with a sub-sample of 50 black schools showed that Per Cent of Females Separated did not contribute independently in predicting achievement levels in these schools. The results are discussed in terms of the possibility that concentrations of urban poverty signify social and institutional disorganization over and beyond the effects of poverty per se. (Author)

FINAL REPORT

Project No. 1-G-012

Grant No. OEG-7-71-0006(509)

An Exploration of the Use of Socioeconomic Census Data to Predict Achievement and Evaluate the Effects of Concentrated Urban Poverty Among Elementary Schools in a Big City

Daniel U. Levine, Robert S. Stephenson,
and Kenneth R. Mares

University of Missouri - Kansas City
School of Education
Center for the Study of Metropolitan Problems in Education
Kansas City, Missouri

February 1, 1973

The research reported herein was performed pursuant to a grant with the Office of Education, U. S Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Regional Research Program

AUTHORS' ABSTRACT

The major purpose of this study was to determine whether data from the 1970 Census could be used to account for at least 70 per cent of the variance in grade-level achievement scores in the elementary schools of a big city, without including variables directly denoting race or ethnicity. This goal was attained when it was found that a regression equation using four variables from the Third Count (block statistics) of the Census accounted for 75 per cent of the variance in sixth-grade achievement scores in a sample of 122 elementary schools in Chicago. The four variables (Per Cent of Females Separated; Per Cent of Families Which Lack 1 or More Plumbing Facilities; Per Cent of Units with 6 Persons or More; Per Cent Owner Occupied Units) also provided the basis for exploratory analysis with a number of additional variables. Separate analysis with a sub-sample of 50 black schools showed that Per Cent of Females Separated did not contribute independently in predicting achievement levels in these schools. The results are discussed in terms of the possibility that concentrations of urban poverty signify social and institutional disorganization over and beyond the effects of poverty *per se*.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I. BACKGROUND AND PURPOSE

In June of 1969, Walter I. Garms and Mark C. Smith completed a study for the New York State Educational Conference Board in which they studied achievement in 45 schools in a stratified sample of differing size cities from various parts of New York state.[1]

Reviewing previous research, they pointed out that socioeconomic background, home environment, and related measures have been found to correlate more highly with achievement than do school-related variables. In addition to showing that socioeconomic status of a student body is a relatively good predictor of achievement, Garms and Smith argued, this finding also suggests that socioeconomic and related variables can indicate whether students in a given school are particularly disadvantaged educationally and thus can serve as a measure of educational need in state aid formulas. Thus the purpose of their study was "to develop a rational and practical measure of educational need and to suggest ways in which such a measure might be incorporated into the method of distributing state financial aid."

After considering a number of possible variables, Garms and Smith decided to work with data on the following student background measures: racial and ethnic background; broken homes; welfare; parental education; overcrowded housing; and mobility. (Several of those variables were measured with two or more alternative sub-variables.) Next they obtained data on achievement in reading and arithmetic and appropriate background data on twenty fourth-grade students in each school in their sample, and then used multiple regression techniques to determine which sets of variables provided the most powerful and economical prediction of achievement levels in those schools. Based on a list of criteria for acceptability of regression equations plus an analysis of residuals, they chose three equations which predicted 70 to 71 percent of the variance as their best predictor equations. Each of the three equations contained the variables Percent

---

[1] Walter I. Garms and Mark C. Smith, "Development of a Measure of Education Need and Its Use in a State School Support Formula." Albany, New York: New York State Educational Conference Board. June 1969. The results of this study also are reported in Walter I. Garms and Mark C. Smith, "Educational Need and its Application to State School Finance," The Journal of Human Resources, v. 5, no. 3 (Summer 1970), 304-317 and in James A. Kelly, "Resource Allocation and Educational Need," Education and Urban Society, v. II, no. 3 (May, 1970), 251-276.

Negro; Percent Puerto Rican; and Average Years of Schooling of Father Where Present, Otherwise Mother. In addition, two equations contained the variable Broken Homes and two contained the variable Mobility. The authors concluded by recommending that "a year be devoted to a required collection and processing of data for all schools in the state, and that a prediction equation be developed based on that solid data base . . . A state aid formula incorporating this prediction equation would be put into use in the following year."

Although Garms and Smith's study constituted an important contribution to thinking and research in educational finance, we believed it raised certain issues and difficulties which required additional study. In particular, we felt that there would be a number of political and administrative problems in using data incorporating measures of race and ethnicity into state school aid formulas. Moreover, several previous studies on determinants of achievement have indicated that race and ethnicity add little to the prediction of achievement after suitable controls are imposed for socioeconomic background and home environment. In addition, it was possible that the variables Percent Negro and Percent Puerto Rican might very well be serving as proxy measures for contrasts between conditions in high density, poverty neighborhoods in large urban centers on the one hand as compared with conditions in less socially disorganized communities on the other. That is, it was possible that the schools with high proportions of Negro or Puerto Rican pupils would tend to be the schools in Garms and Smith's sample which were located in high density, very low-income poverty neighborhoods in cities in New York State. Since there is reason to believe that conditions in high density, poverty neighborhoods constitute a special set of circumstances in which "threshold" limits involving the operation of social controls have been exceeded, and since social disorganization beyond given threshold levels almost certainly leads to rapid deterioration in social institutions such as the schools, it seemed possible that the equations developed in Garms and Smith's linear regression analysis were picking up effects of dense, urban poverty rather than race or ethnicity per se.

Accordingly, we obtained access to Garms and Smith's data and carried out a re-analysis to determine whether a non-linear analysis omitting the variables Percent Negro and Percent Puerto Rican could account for as much of the variance as had been predicted in the original study. This was done by including squared and cubed terms and interaction terms for most of the remaining variables, in

order to test whether the introduction of these terms would reveal curvilinear changes in the relationship between student background variables and achievement in extremely low-status (or extremely high-status) schools in the sample.

The results of this analysis were sufficiently encouraging to warrant additional research for the specific purpose of exploring the possibility of using socioeconomic variables (but not race or ethnicity) for state aid formulas or the general purpose of learning more about the effects of concentrated urban poverty on school achievement. One of our equations, for example, showed that the five variables welfare, welfare cubed, father's schooling, welfare x broken homes, and broken homes x father's schooling accounted for 71% of the variance in arithmetic achievement among the schools in Garms' and Smith's sample. For reading achievement, 59% of the variance among schools was accounted for by the six variables welfare, welfare cubed, foreign language spoken in the home, broken homes cubed, mobility, and broken homes x mobility. In the full stepwise equation, the cubed terms for welfare and broken homes were more powerful predictors of achievement than were the squared terms for these variables; this finding suggests that social disorganization as measured by percent on welfare and percent from broken homes may have a particularly detrimental impact on school achievement in neighborhoods where these percentages are extremely high.

This re-analysis, however, could not be interpreted as being more than suggestive of the need for further research. Due to problems in the original study, particularly skewedness in the data and difficulties in the measurement of variables, neither Garms and Smith's linear analysis nor our own nonlinear analysis can be viewed with much confidence in the absence of more refined investigations with other samples. In the present study of the relation between census data and school achievement in one big city, accordingly, our major purpose is to determine whether sets of student background variables excluding race and ethnicity can account for as much of the variance in achievement (approximately 70%) as Garms and Smith explained in their New York State Study.

More particularly, it was hypothesized that variables and methodology which took dense urban poverty into account either by describing conditions in a big city poverty neighborhood or by non-linear treatment in a multiple regression equation would make it possible to achieve this objective. It was decided to collect data in Chicago, primarily because school-by-school ability and

achievement data were available to correlate with 1970 Census data on neighborhoods and communities and because Chicago is a large city with large low-income neighborhoods in which the effects of concentrated poverty should show up relatively clearly. One reason for obtaining a sample of schools from one big city school district is to control as much as possible for district-level variations in test administration, promotion policies, and other variables that might obscure relationships in multi-district comparisons. The achievement and student background variables chosen for inclusion in this study are described in the following chapter on Methods and Procedures.

Relation to Previous Research

One of the basic issues involved in this study is whether race or ethnicity is associated with achievement after account is taken of differences in socioeconomic status and family background. There is, of course, an enormous literature on numerous aspects of this question. As one would expect regarding a body of data so large and voluminous, various researchers have arrived at opposing conclusions and the research as a whole is far from conclusive. For our purposes here it is not necessary to provide a comprehensive review of the literature; instead we will cite only a few of the major studies and viewpoints on the issue.

To begin with, it is well established that indices of socioeconomic status (e.g., father's occupation, father's education, family income) typically correlate with school achievement at .3 to .4 even among samples of students with a relatively restricted range of social class backgrounds.[2] Furthermore, researchers who have obtained information on social-class-related home and family environment variables (e.g., intellectual expectations for the child; opportunities for learning in the home) more directly related to school achievement than social status *per se* have reported correlations as high as .8 between home environment and achievement.[3] Peterson and DeBord, for example, studied home environment variables among low-status students in Nashville and found that a carefully-selected set of five of the variables correlated (multiple r) at .71 with achievement among black subjects and a set of six variables was correlated at the same

---

[2] Lavin, D. E., The Prediction of Academic Performance (New York: Science Editions, 1967).

[3] Robin H. Farquhar, "Home Influences on Achievement and Intelligence: An Essay Review," Administrator's Notebook, v. 13, no. 5 (January 1965).

level with achievement among white subjects.[4] Since home environment data thus have explained as much as two-thirds of the variation in achievement among students in several carefully-designed studies, it is warranted to view social and cultural background as a major if not predominant determinant of achievement in the schools.

The well-known Equality of Educational Opportunity study conducted by Coleman, Campbell, et al. reported that the social-class background of individual students accounted for more than 80% of the variation in school achievement and that socioeconomic integration accounted for most of the increment in achievement found among minority students in racially integrated schools.[5] That is, only a negligible part of the achievement differences of minority students in segregated and integrated schools could be attributed to racial integration after the effects of socioeconomic integration had been accounted for. Using alternate approaches in analyzing data from the equal opportunity study, the U. S. Civil Rights Commission issued a report concluding that racial integration in the schools in certain circumstances was a major factor in improving the performance of minority students.[6] However, Nichols took issue with this interpretation after arguing that possible differences between the families of similar-status minority students in segregated and integrated communities had not received sufficient attention in the Commission analysis.[7] Indeed, a study by Wilson which was done for the Commission and reported in an appendix to its main report and which took account

---

[4] Richard A. Peterson and Larry DeBord, Educational Supportiveness of the Home and Academic Performance of Disadvantaged Boys. Nashville, Tennessee: George Peabody College for Teachers, 1966.

[5] James S. Coleman, et al., Equality of Educational Opportunity (Washington, D. C.: U. S. Government Printing Office, 1966). Most of the controversy over the equal opportunity study has centered on the possibility that school effects may have been underestimated rather than on the question of socioeconomic versus racial determinants of achievement.

[6] Racial Isolation in the Public Schools, Volume I. A Report of the United States Commission on Civil Rights (Washington, D. C.: U. S. Government Printing Office, 1967).

[7] Robert Nichols, Review of Racial Isolation in the Public Schools, American Educational Research Journal, v. 5, no. 4 (November 1968), 700-707; see also "The Coleman Report: An Inadequate Study but the Best We Have So Far," Phi Delta Kappan, June 1967, 527-530, reprinted from Science, December 9, 1966.

of family differences by utilizing data on first grade ability concluded that the socioeconomic status of individuals and the socioeconomic composition of schools were much better predictors of achievement than race. This conclusion was further supported by the results of an exhaustive review of the available research literature carried out by Robert O'Reilly of the New York State Department of Education.[9] Thus the preponderant conclusion suggested by these major studies and research reviews is that one can predict a large amount of the variation in school achievement with information on the socioeconomic background of individual students and the socioeconomic composition of their schools but none whatever on the racial or ethnic composition of the schools.

Even granted that a substantial correlation exists between socioeconomic and family background (exclusive of race) on the one hand and school achievement on the other, however, it is still possible that race is a contributing variable which might account for much of the remaining unexplained variation. Jensen, in fact, has made just this argument in a major paper arguing that blacks as a group are genetically inferior on conceptual tasks which are emphasized in widely-used ability tests and which presumably are necessary for success in school, especially in the later grades.[10] Jensen's paper provoked a storm of controversy which has not subsided. Following its publication, a number of prominent authorities in psychology, genetics, education, and related fields criticised and took exception to many of his methods and conclusions.[11] More recently, Rohwer presented evidence which called into question Jensen's fundamental distinction between "conceptual" and "associative" learning.[12] While we cannot

---

[8] Alan B. Wilson, The Consequences of Racial Segregation Berkeley, California: Glendessary, 1966).

[9] Robert O'Reilly, et al., Racial and Social Class Isolation in the Schools (Albany: New York Board of Regents, 1970).

[10] Arthur R. Jensen, "How Much Can We Boost I.Q. and Scholastic Achievement?", Harvard Educational Review, v. 39 (1969), 1-123.

[11] See particularly the Spring, 1969 and Summer, 1969 issues of the Harvard Educational Review (v. 39) which were devoted in part to responses to Jensen's article in the Winter issue.

[12] William D. Rohwer, Jr., "Learning, Race, and School Success," Review of Educational Research, v. 41, no. 3 (June, 1971), 191-210.

here review the many criticisms which have been made of Jensen's position or the defenses he and others have offered in rejoinder, we can report that the preponderance of expert opinion appears to agree with the following summary which geneticist I. I. Gottesman provided to a chapter on the "Giogenetics of Race and Class" published in a 1968 volume co-edited, interestingly, by Jensen:

> . . . At the present time Negro and white differences in general intelligence in the United States appear to be primarily associated with differences in environmental advantages. Social-class differences in general intelligence in stratified, open-class societies appear to be moving in a direction where such differences will have an appreciable genetic component . . . /but/ So long as persons at the lower end of the I.Q. distribution are at a reproductive disadvantage, that is, less fit, there will be positive selection for this prized human trait.[13]

Only a few empirical studies have focused directly on the question of whether racial differences persist after social class and/or family environment are controlled. Several studies, mostly small in scale and limited in the types of abilities and achievement tested, have reported no difference between subjects of differing racial groups. Grimmett, for example, reported that there were no differences in problem-solving behavior among low-status "Mexican-American, Black-American, Indian-American . . . /and/ Appalachian Caucasian-American" third and sixth graders.[14] Musgrove and Lawson, similarly, found no difference in performance on two ability tests between samples of low-status white and black children, though they did find differences on an information test.[15] Stodolsky and Lesser, however, and some other researchers have reported finding differences

---

[13] I. I. Gottesman, "Biogenetics of Race and Class," in Martin Deutsch, Irwin Katz, and Arthur R. Jensen (eds.) Social Class, Race, and Psychological Development (New York: Holt, Rinehart, and Winston, 1968), p. 46.

[14] Sadie Grimmett, The Influence of Ethnicity and Age on Solving Twenty Questions. Nashville, Tennessee: George Peabody College for Teachers, n.d. in Martin Deutsch, Irwin Katz, and Arthur R. Jensen (eds.), op. cit.

[15] Walter L. Musgrove and John R. Lawson, "A Comparison of Lower Class Negro and White Children on Three Standardized Tests," The Journal of Negro Education, v. 40, no. 1 (Winter 1970), 53-55.

in various types of abilities between youngsters from several racial and ethnic groups, even after controlling for social status.[16]

Several studies which investigated both racial and socio-economic variables are particularly intriguing because of their relative excellence in design and because of the incompatibility of their conclusions. Deutsch and Brown examined intelligence test scores of black and white subjects from three social class levels and concluded that lower-class status had a similar effect on the two groups but that higher status blacks and whites differed in performance on ability tests with the former scoring lower than the latter.[17] They did not, however, perceive a genetic basis for this difference but instead hypothesized that blacks experienced relatively greater obstacles to cultural and social participation as they rose in social status. A somewhat comparable conclusion was reported by Bachman in 1970 as part of a major study of the impact of background variables on the performance of tenth-grade males. Based on a sample of 2213 subjects in 87 public high schools throughout the United States, Bachman reported that black students most of whom presumably were low-status in southern segregated schools performed much more poorly than other subjects after account was taken of social status and family background but that black respondents in integrated schools have scores similar to whites when social status was controlled.[18]

In a study published in 1968, Tulkin collected data on fifth and sixth graders and reached a conclusion precisely opposite to that of Deutsch and Brown. After equating for social status and family environment, he found differences in performance on ability and achievement tests between lower status but not higher status

---

[16] Gerald Lesser and Susan S. Stodolsky, "Learning Patterns in the Disadvantaged," Harvard Educational Review, v. 37 (1967), 546-593.

[17] Martin Deutsch and Burton Brown, "Social Influences in Negro-White Intelligence Differences," Journal of Social Issues, 1964 (20), 24-35.

[18] Jerald G. Bachman, Youth in Transition. Volume II. The Impact of Family Background and Intelligence on Tenth-Grade Boys (Ann Arbor, Michigan: The University of Michigan Survey Research Center, 1970). It is possible that black subjects in integrated schools had more supportive home environments, on the average, than their counterparts in segregated schools.

white and black subjects.[19] Tulkin suggested that low-status blacks are at a lower economic level and have a higher incidence of problems such as prematurity than similar-status whites and that differences of this type may account for the ability and achievement differences he found between low status black and white students.

One of the reasons why some researchers have found a difference in performance between high status but not low status blacks and whites while others have found the reverse pattern may be due to the extreme difficulty of devising comparable social status scales for differing racial and ethnic groups. That is, social status designations for a given group may have quite a different meaning when applied to another group and a given method for assessing social status may be more meaningful for one group than another. Among whites, for example, the occupation of postal worker had much less status than it did, at least until recently, for blacks. Since black postal workers were likely to have higher social status in their communities than did white postal workers, it is also probable that the two groups differed more widely in child-raising practices, press for school achievement, and other home and family variables than would be true if occupation represented similar levels of social status for the two groups. In view of the difficulties encountered in designating comparable social class placements for whites and blacks, it is not surprising that researchers studying social-class related phenomena sometimes reach contradictory conclusions.

In addition, there is also evidence that specific home and school variables have differential importance for differing racial and ethnic groups.[20] Peterson and

---

[19] Steven R. Tulkin, "Race, Class, Family, and School Achievement," Journal of Personality and Social Psychology, v. 9, no. 1, (1968), 31-37.

[20] On differential effects of home variables, see Richard A. Peterson and Larry DeBord, op. cit., and Lesser and Stodolsky, op. cit.
On differential effects of school variables, see James S. Coleman, et al., op. cit., and Stephan Michelson "The Association of Teacher Resourcefulness with Children's Characteristics," in Do Teachers Make A Difference? A Report on Recent Research on Pupil Achievement (Washington, D. C.: U. S. Government Printing Office, 1970). Although Michelson concluded that certain school variables appear to have differential effects on black and white students,

DeBord, for example, found that church attendance of the mother was positively associated with achievement among low-status white students but negatively associated with achievement among low-status black students. Because of these differentials, much of what a researcher finds in studying home and school influences on achievement will depend on his choice of variables and whether his analysis allows for the emergence of these differential effects.

The major conclusion which can be drawn from this part of our review of previous research is that social status and family variables can allow one to predict two-thirds or more of the variation in scores on ability and achievement tests (some of the remaining variance is ipso facto unpredictable due to test unreliability). Race and ethnicity per se may or may not add significantly to predictive power after cultural and social influences associated both with performance and with race and ethnicity are taken into account, but in any case race and ethnicity generally account for only a small fraction of the variance in performance on ability and achievement tests after social status and family environment are first taken into account. How, then, explain Garms and Smith's finding that race and ethnicity predicted a large share of variation in achievement even after accounting for social status through multiple regression analysis?

First, Garms and Smith had no good direct measure of home or family environment; thus much of the variance accounted for by race and ethnicity in their equations actually may be reflecting differences in family and home environment among racial and ethnic groups in their sample. Second, it is highly probable that low-status subjects from racial and ethnic minority groups (Negroes and Puerto Ricans) in their sample were more frequently concentrated in urban poverty ghettoes than were their other subjects. Since there is reason

---

he reported that "school characteristics which affect whites, particularly middle class whites, are different from those characteristics which affect blacks and lower class whites." This finding is more compatible with our own hypothesis concerning the relation between concentrated urban poverty and school achievement than it is with Michelson's conclusion concerning differential school effects. (Presumably, lower class whites are similar to blacks in that they tend to be concentrated in urban slums or near-slums.)

to believe that urban density, particularly concentrations of urban poverty, may have detrimental effects on the functioning of social institutions such as the family and the school (see below), it is possible that racial and ethnic variables in the Garms and Smith study were acting at least in part as proxy variables for big-city urban poverty and urban social disorganization. Considerations of this nature suggested that variables and methodology indicative of concentrated poverty in a big city (but excluding politically-troublesome variables describing race and ethnicity) might enable us to account for as much of the variation in performance (70-71%) as Garms and Smith had explained.

Very little research has been carried out explicitly to determine whether concentrated urban poverty causes institutional dysfunctioning or otherwise generates social disorganization to a greater extent than one might predict given the well-established association between these conditions and social status. However, biologists and ethologists have left little room to doubt that density frequently has pathological effects in non-human populations, and many social as well as natural scientists have expressed the opinion that density frequently has similar effects in human populations as well.[21] Distinguished social scientists such as Edward T. Hall have argued that density has debilitating effects on the inhabitants of urban poverty neighborhoods and have identified some of the social and psychological processes which probably function to generate social disorganization in homogeneous low-status environments.[22] Indeed, even "common sense" observation on this topic has become so undeniable that a federal court has ruled it illegal for the Chicago Housing Authority to build housing developments which would allow more than 15% of the population of a census tract to be low-status residents of public housing projects.[23]

---

[21] Leo Levy and Harold M. Visotsky, "The Quality of Urban Life: An Analysis from the Perspective of Mental Health," in Urban America: Goals and Problems (Washington, D. C.: U. S. Government Printing Office, pp. 104-105.

[22] Edward T. Hall, The Hidden Dimension (Garden City, N.Y.: Doubleday, 1966).

[23] J. S. Fuerst, "First Priority: Replace Chicago Housing Administration," Focus/Midwest, v. 7, no. 48, 7-10.

Only a few studies have provided direct empirical evidence bearing on the proposition that concentrations of urban poverty generate social disorganization and institutional dysfunctioning. A study conducted in Washington, D. C. by James T. Dailey concluded that effective school reading programs were much more difficult to operate when 90% of the students enrolled were below national norms (as sometimes is true in classes in inner city schools) than when the percentage was above 90%.[24] Boyle reviewed studies prior to 1966 on the relation between students' aspirations and school social status and concluded that segregation by social class has a detrimental effect on aspirations of youth in large urban areas,[25] but this effect is not so evident in small urban areas (which presumably have lower densities on the average); studies on this issue conducted since 1966[26] are fully compatible with Boyle's interpretation. While we found no other data which directly supported the proposition, neither did we find any empirical research which contradicted it or called it into question. Thus the present study provides one of the few tests explicitly designed to assess the proposition that concentrated urban poverty has detrimental effects on low-status individuals over and beyond the educational and other disadvantages associated with low-status background per se.

Importance and Unique Contribution of the Study

In addition to possibly clarifying the dynamics of urban development by determining whether the public schools are adversely affected when concentrations of poverty exceed a hypothetical threshold limit, this study potentially may make a unique contribution to research and thinking in urban education by showing that easily-available social and economic data exclusive of race and ethnicity can account

---

[24] John T. Dailey, "Evaluation of the Contribution of Special Programs in the Washington, D. C. Schools to the Prediction and Prevention of Delinquency," Washington, D. C.: George Washington University, 1966, p. 39.

[25] Richard P. Boyle, "The Effect of the High School on Students' Aspirations," American Journal of Sociology, v. 71, 628-39.

[26] Daniel U. Levine, Edna S. Mitchell, and Robert J. Havighurst, Opportunities for Higher Education in a Metropolitan Area. (Bloomington, Indiana: Phi Delta Kappan, 1971).

for as much as 70% of the variance in school achievement. One may ask how such a contribution would be unique in view of the fact that Dave-Wolf and others already have demonstrated that socioeconomic and family background data can account for at least this much of the variation in school achievement. The answer is that most previous studies accounting for this much of the variance in school achievement either have depended on laboriously-collected interview data from necessarily small samples of parents or have had to include variables describing race and ethnicity (e.g., Garms and Smith). If we can establish, by way of contrast, that data easily retrievable from the federal census can account for two-thirds or more of the variation in school achievement in a large urban area, it would be much more administratively and politically feasible to rewrite state school-aid formulas or otherwise take action to reduce educational inequalities attributable to social and economic deprivation than is presently possible given political and economic realities in state government decision-making for education.

CHAPTER II. METHODS AND PROCEDURES

To determine whether census data, particularly data bearing on concentrated urban poverty, could explain as much as 70% of the variance in school achievement in a big city school district, we had to obtain data from a sample of schools in a city in which appropriate information on the schools was available for analysis in conjunction with data on the communities served by the schools in the sample. Several considerations made Chicago a good choice in which to carry out the study.[1] Accordingly data on schools and neighborhoods in Chicago were obtained and analyzed using the methods and procedures described in this chapter.[2]

Choice of schools

After obtaining a map showing the location of elementary schools in Chicago, we selected a geographically- and socially-diverse sample of approximately 185 elementary schools which included grades one through six.[3] Since data were available on the racial and ethnic composition of Chicago elementary schools, and since one of the investigators was personally familiar with schools in Chicago, it was not difficult to make sure that the sample included schools which varied widely

---

[1] These considerations included the following: school-by-school ability and achievement scores were a matter of public record in Chicago; the city school system includes a sufficiently large and diverse number of elementary schools (more than 500) to facilitate drawing a sample with whatever characteristics are desired; and the city itself is diverse in socioeconomic characteristics but unquestionably has many dense poverty neighborhoods.

[2] We would like to acknowledge and express appreciation for the kind cooperation extended by many Chicago school officials, particularly the following:

Irving Brauer - Office of Operations Analysis
Dr. Elmer Casey - Director of Pupil Evaluation and Studies
Robert Kelley - Director of Administrative Research
Thomas Teraji - Director of Attendance Area Studies
Frank Ward - Division of Pupil Evaluation and Studies

[3] There is much variation in organizational patterns in the Chicago Public Schools. In addition to pre-school centers there are a number of primary-grade centers and K-8 schools, as well as other variations. Schools included in the preliminary sample ranged from K-6 to K-8.

in racial and ethnic composition. A special effort also was made to include predominantly black schools in relatively high status neighborhoods and predominantly white schools in relatively low status neighborhoods as well as particularly high and low achieving schools of whatever composition, in order to facilitate later analysis examining the relationships between social status, race and ethnicity, and school achievement.

Next, knowledgeable school district officials were consulted in order to determine whether special circumstances existed which might make it inappropriate to include any of the schools in the preliminary sample. Approximately fifteen schools were eliminated at this stage either because they did not serve surrounding neighborhood populations to the same extent as do most elementary schools in Chicago[4] or because critical problems usually involving school-community relations had kept them in constant turmoil during the year achievement data had been collected which were to be used in the study.

After the preliminary sample had been constituted in this way, official maps showing the attendance boundaries of elementary schools in Chicago were obtained. Using metropolitan maps of Chicago produced by the U. S. Census Bureau, a list was made showing the block groups and census tracts within each school's attendance area. At this stage in the selection process, only block groups which were entirely within and census tracts which visual inspection indicated were at least 80% within a given school's attendance area were included in the list of block groups and tracts for that school. Using these lists, it now was possible to extract information on block groups contained in the Third Count of the 1970 Census and information on census tracts contained in the Fourth Count.

However, to make sure that we would obtain as close as possible a match between student achievement data on the one hand and socioeconomic data on the neighborhoods in which students lived on the other, three additional steps were taken in selecting the final sample of schools. First, "spot" maps showing the actual block location of students enrolled in the schools in the sample as of the spring of 1970 were obtained and used to superimpose "effective" school boundaries

[4] Schools eliminated on this basis included several involved in a busing program to relieve overcrowding on the west side and several others which were special education centers.

(i.e., the area from which schools actually draw most of their students as contrasted with the official boundary lines) upon the larger metropolitan map showing the location of census tracts and block groups.[5] Second, 145 schools were selected from within the preliminary sample which visual inspection indicated had the most satisfactory "fit" between effective attendance area and census tract boundaries. Third, block group data from the Third Count (see below) were obtained for these schools and a comparison was made between data on demographic variables available from the Census and data on student population characteristics during the same year (1970). Comparing the two types of data, 21 schools were found in which the percentage of white students in the school was 15 percentage points or more smaller than the comparable percentage in the attendance area. On the assumption that special circumstances existed in these communities which made it inadvisable to view the school population as representative of its neighborhood (for which census data would be related to school achievement data),[6] these schools were eliminated. In addition, two schools were eliminated because there were less than 1,000 people on the respective blocks for which we extracted third count data for their neighborhoods. Thus, 122 schools remained to comprise the basic sample. Afterwards, lists were made showing the block groups and census tracts which fell within the effective attendance area for each of these schools.

---

[5] In determining "effective" school attendance areas, several decision rules were developed after preliminary study of attendance patterns at a few schools. First, we used data only on students in grades 1 through 6. Second, we generally did not include a block unless at least two students were enrolled in the school or the next block further out had at least three students enrolled. Third, slight departures from these criteria were tolerated when doing so made the effective attendance area of a school much more compact than would have been the case had the criteria been universally applied.

[6] Special circumstances which might account for this discrepancy between the school data and the community data include: a particularly high percentage of white students may attend private and parochial schools; the white population may be relatively old and without children in the home, as often happens in neighborhoods undergoing racial transition; or residential turnover may be so rapid that school and census data collected even a few months apart may not be easily comparable.

Third Count Census Data

The study described in this report utilized the Third Count of the decennial census conducted in the City of Chicago by the U. S. Bureau of the Census in the spring of 1970. The third count is the only tabulation that provides data by blocks - the smallest unit for which data are collected and released.

Data were extracted from computer tapes at the Demographic Data Center, Technical Services Division, at the Columbia campus of the University of Missouri. The computer programs developed for this purpose used lists of the blocks (by census tracts) we had selected to represent school neighborhoods in accordance with the procedures described above. The computer printouts showed percentages based on the total number of persons, houses, etc. (depending on the variable) in the blocks chosen for each particular school. In some cases variables were defined in the same way as they are presented in the census; in other cases, we designated our own variables based on raw data from several census categories. For this reason much of our data cannot be derived directly from U. S. Census Bureau Reports (print, microfiche, or tape) without additional computations similar to those carried out in our computer programs.

Following this procedure, we selected and constructed eighteen variables describing neighborhood conditions in our 122-school sample. Several of these variables seemed particularly promising as indicators of concentrated urban poverty (e.g., Variable 17, below). The 17 variables, together with the abbreviated names we used for each variable are shown in Table 1.

As mentioned in the preceding chapter, one of the purposes of this study was to determine whether taking account of concentrations in urban poverty neighborhoods could substantially improve predictions of school achievement levels based on demographic data from the U. S. Bureau of the Census. Stated differently, this objective hypothesizes that the relation between school achievement data and socioeconomic data on pupils in urban areas is not entirely linear. For example, disadvantaged students in neighborhoods of concentrated poverty (such as inner city slums) may achieve less than one would predict given the performance of similar-status students in working-class or middle-class neighborhoods. More generally, this hypothesis presumes that concentrated poverty areas frequently may be characterized by such serious social disorganization that the schools and other social institutions function extremely poorly even as compared with similar institutions in neighborhoods which are only a little wealthier.

TABLE 1

Third Count Variables and Abbreviated Titles

Variable	Title	Mean	Standard Deviation
1. Per Cent Negro	PCNEG	41.38	47.80
2. Per Cent Negro and Other	PCNOTH	42.92	46.91
3. Per Cent of Females Separated	PCFSEP	6.31	5.76
4. Per Cent Owner Occupied of Total Occupied Units	PCOWN	36.65	27.81
5. Per Cent of Housing Units Which Are One-Unit Structures	PCONE	27.32	29.36
6. Per Cent of Units with 6 Persons or More	PC6PER	13.50	8.16
7. Per Cent of Occupied Housing Units 1.51 or More Persons Per Room	PC151	3.66	3.47
8. Per Cent of Negro Renter Occupied Units With 1.51 or More Persons Per Room	PCNR151	2.64	3.95
9. Per Cent of Housing Units Not With Flush Toilets For Household Only	PCNOTOY	3.12	4.51
10. Per Cent of Occupied Housing Units Worth Less Than $9999	PCLT99	64.70	36.17
11. Per Cent of Renter-Units with Rent Less Than $59	PC59	23.57	28.82
12. Per Cent of Renter-Units with Rent Less Than $99	PC99	56.59	27.31
13. Per Cent of Housing Units Vacant Six Months or More	PCVAC6	2.41	3.21
14. Per Cent of Occupied Units with Female Head Families	PCFHED	15.70	9.91
15. Per Cent of Families Which Lack 1 or More Plumbing Facilities	PCLACK	2.08	2.06
16. Per Cent of Occupied Units with Negro Occupants Lacking 1 or More Plumbing Facilities and Having 1.01 or More Persons Per Room	PCNOLA	.14	.30
17. Per Cent of Population in Units Lacking 1 or More Plumbing Facilities and with 1.01 or More Persons Per Room	PCPLACK	.72	.82

The main approach by which we tried to take account of this kind of possible non-linearity was by including interaction terms (i.e., multiplying one possible indicator of concentrated poverty by another) and squared or cubed terms (for such indicators) in statistical analyses using pupil achievement data as the dependent variable and socioeconomic (i.e., neighborhood) and/or school characteristics data as the independent variables. Thirteen additional neighborhood variables consisting of such terms chosen for this purpose are shown in Table 2.

TABLE 2

Squared, Cubed, and Interaction Variables Derived From Third Count Data

	Variable*	Title+
1.	PC99 X PCNOLA	127 X 31 (61)
2.	PC99 X PCPLACK	127 X 36 (62)
3.	PC99 X PCNR151	127 X 16B (63)
4.	PCNR151 X PCLACK	116B X 30 (64)
5.	PCNR151 X PC59	116B X 23A (65)
6.	PC151 Squared	50 X 50 (67)
7.	PC151 Cubed	Cube 50 (68)
8.	PCFHED Squared	57 X 57 (69)
9.	PCFHED Cubed	Cube 57 (70)
10.	PCPLACK Squared	60 X 60 (71)
11.	PCPLACK Cubed	Cube 60 (72)
12.	PC151 X PCFHED	50 X 57 (73)
13.	PCFSEP X PC59	45 X 54 (74)

*Variables listed in this column are explained in Table 1.
+Terminology in this column was chosen for its convenience in data processing and has no inherent meaning.

School Characteristics and Achievement Data

Data on the characteristics of and achievement in schools in the sample were taken from several reports released by the Chicago Board of Education. For most variables, we used information collected as close in time as possible to the date of the Census (September, 1970); in a few cases, however, we extracted additional data collected at a more distant time if it appeared likely that this information might be useful in subsequent analysis. Variables constructed from these data in order to describe school characteristics and achievement in the 122-school sample are shown in Table 3, together with their titles, mean scores, and standard deviations.[7]

[7] To save space, a number of variables which were extracted but not used in subsequent analysis due to redundancy and/or data processing limitations are not included in Table 3 or described explicitly elsewhere in this report.

TABLE 3

School Characteristics and Achievement Variables*
(N = 122)

Variable+	Title	Mean	Standard Deviation
1. Per Cent Caucasian (Sept. 1970)	PCCAUC	43.89	41.89
2. Per Cent African or Negroid Origin (Sept.1970)	PCAFRI	42.93	48.30
3. Per Cent Mexican Origin (Sept. 1970)	PCMEX	5.48	14.38
4. Per Cent Puerto Rican Origin (Sept. 1970)	PCPUER	5.17	11.17
5. Per Cent Spanish Background (Mex, PR, & Cuban, Sept. 1970)	PCSPAN	11.35	20.56
6. Per Cent English Second Language (Spring, 1970)	PCENGS	8.47	16.45
7. Learning Ability Percentile, 1st Grade Q2 (Jan. 1970)	SLAB1P	56.79	19.63
8. Learning Ability Percentile, 4th Grade Q2 (Apr. 1970)	SLAB4P	44.17	18.39
9. Learning Ability Percentile, 6th Grade Q2 (Mar. 1970)	SLAB6P	43.34	20.93
10. Learning Ability Percentile, 1st Grade Q3 (Jan. 1970)	SLA1Q3	75.84	14.71
11. Learning Ability Percentile, 4th Grade Q3 (Apr. 1970)	SLA4Q3	64.89	15.10

*This list includes variables describing student ability scores.

+Unless otherwise noted, school characteristics data are for grades 1-8 or 1-6, depending on each school's organization. Tests for which data were obtained are as follows: Readiness at entrance to first grade: Metropolitan Readiness Test; First grade: Kuhlmann-Anderson Intelligence Test; Fourth Grade: Kuhlmann-Anderson and Metropolitan Achievement Battery, Form B Reading; Sixth Grade: California Short-form Test of Mental Maturity, 1957-S Form and Metropolitan Achievement Test, Reading. Q2 scores represent the median in a given sample or sub-sample, and Q3 scores represent the point that divides the top quarter from the remaining three quarters. On nationally-standardized tests, the Q2 national percentile "norm" is 50 and the Q3 norm is 75. Scores listed as "4th" Grade actually represent data obtained at the end of the primary-intermediate transition (3rd) grade.

TABLE 3 (Continued)

Variable+	Title	Mean	Standard Deviation
12. Learning Ability Percentile, 6th Grade Q3 (Mar. 1970)	SLA6Q3	64.92	19.01
13. Pupils per Teacher (Spr. 1970)	TAPP	32.89	2.45
14. Per Pupil Classroom Teacher Expenditure (Oct. 1970)	CTEXPP	363.35	41.53
15. Per Pupil Staffing Cost (Total School, Spr. 1970)	PPSTAFTL	413.05	55.51
16. Total Expenditure Per Pupil (Oct. 1970)	TCEXPP	449.62	58.49
17. Per Cent Caucasian Teachers (Oct. 1970)	PCCAUT	68.43	29.40
18. Date of Construction	DATE	1927.64	25.44
19. Per Cent Faculty with Six Years or More Exper. (Spr. 1970)	FAC6PL	52.70	22.41
20. Per Cent Entering First Year Ready (Sept.1969)	EFYRRD	57.89	23.98
21. Reading Percentile, 4th Grade Q2 (Apr. 1970)	RQ24TH	43.08	17.86
22. Reading Percentile, 6th Grade Q2 (Mar. 1970)	RQ26TH	36.00	18.70
23. Reading Percentile, 4th Grade Q3 (Apr. 1970)	RQ34th	59.18	17.26
24. Reading Percentile, 6th Grade Q3 (Mar. 1970)	RQ36TH	54.25	18.95
25. Classroom ratio including specialists and supervisors	SACRIS	33.77	1.60

CHAPTER III. RESULTS

The first step in our analysis of the data was to obtain a correlation matrix showing the association between all the neighborhood (census) and school variables in the study.[1] This matrix enabled us to pick out the best combinations of census variables for predicting school achievement scores--that is, the census variables which were most highly correlated with achievement but least inter-correlated among themselves.

Preliminary stepwise regression analyses also were carried out at this point. The regression data indicated that there was considerable stability in the functioning of most of the census variables in the sense that these variables generally entered regression equations with roughly the same weight and direction of influence and approximately in the same order at differing grade levels. This finding is important if one hopes to identify census variables which can be used to obtain a good prediction of school achievement in future replications without repeating all the preliminary work necessary in this exploratory study.

Our preliminary analysis also indicated that errors had occurred in extracting neighborhood variables dealing with the value of housing (variables 10, 11, and 12, Table 1) from the Third Count Census Tape. For this reason, neither these variables nor interaction terms including any of them could be utilized in subsequent analyses reported in this paper. Thus we had to proceed without the most direct measures of community income level to be found in the Third Count. This definitely presented a handicap in our attempt to account for 70 per cent or more of the variance in school achievement scores using data from the Third Count.

Finally, the preliminary analysis showed that a higher percentage of variance in achievement generally could be accounted for at the sixth grade level than at the fourth grade level. (As shown in Table 3, these were the only two grades for which achievement data were available in our sample of schools including grades one through six.) On the one hand, this pattern might have been expected inasmuch as grade six scores in Chicago in 1969 and 1970 might have been less affected by certain compensatory education programs such as Head Start than were scores in the lower grades. In addition, tests administered in upper grades generally tend to be more reliable

[1] Most of the analysis reported in this chapter was carried out on the University of Missouri - IBM 360-65 Computer.

than are those administered in the lower grades. On the other hand, one might predict that sixth grade scores would be more heavily influenced by school variables such as teacher experience than would achievement scores obtained when pupils had been in school for fewer years, thus possibly reducing otherwise high associations between census variables and school achievement. For these reasons it would have been difficult to make a confident a priori prediction about the relative accuracy and usability of test scores at differing grade levels in the Chicago Schools.

At any rate, we did find that census variables were more closely related to achievement scores at the sixth grade than at the fourth grade. Since our goal was to determine or demonstrate the maximum amount of variance (with a target of 70 per cent) in grade-level achievement scores that could be accounted for using a limited amount of easily-obtainable 1970 census data, this chapter deals most often with regression data obtained for the sixth grade. It should be kept in mind, however, that the patterns of relationships almost always were the same at the two grade levels.

Maximum Variance in School Achievement Explained by Census Variables

As noted in the preceding section, the first step in our analysis was to pick out the best combination of census variables to explain (i.e., account for) the maximum amount of variance in the average achievement scores among the sixth grades in our 122-school sample. For this purpose we excluded variables for which regression coefficients were not in the expected direction, that is, where a higher incidence of a poverty-related characteristic was associated with higher rather than lower achievement. (We will discuss several of these variables in a later section of this paper.) It should be noted that several combinations of six or more census variables resulted in equations which accounted for more than 80 per cent of the variance in sixth-grade achievement scores, but we wanted to select an equation that would be stable across grade levels and would include a smaller number of variables which embodied a negative relationship between poverty and achievement. The equation which best met this goal used census data on Per Cent of Females Separated (PCFSEP - Table 1, #3), Per Cent of Families Which Lack One or More Plumbing Facilities (PCLACK - Table 1, #15), Per Cent of Occupied Units with 6 Persons or More (PC6PER - Table 1, #6), and Per Cent Owner Occupied of Total Occupied Units (PCOWN - Table 1, #4) as independent variables and Q3 Reading Achievement (RQ36TH - Table 3, #24) as the dependent variable. Data on the forced order regression equation using census variables in the order given above are shown in Table 4.

TABLE 4. Per Cent of Variance in Sixth Grade Q3 Scores Accounted for by Four Selected Census Variables*

Variable	R	R^2	Increase in R^2	F_1	Regression Coefficient	Standard Error	F_2
PCFSEP	.790	.624	-	-	-.890	.283	9.876 ($<$.005)
PCLACK	.813	.661	.037	13.1 ($<$.001)	-1.283	.500	6.593 ($<$.05)
PC6PER	.830	.690	.028	10.5 ($<$.005)	-.838	.156	28.923 ($<$.001)
PCOWN	.865	.749	.060	27.8 ($<$.001)	.259	.049	27.846 ($<$.001)

*See Tables 1 to 3 for definitions of the variables.

R is the multiple correlation with Q3 sixth grade achievement scores. (However, for the first variable entered, the correlation of course is not multiple.)

R^2 is the square of the multiple correlation and represents the percent of variance accounted for in the independent variable.

F_1 is the F value for the variance added to the equation. The significance level is shown in parentheses.

F_2 is the F value for the weight of the regression coefficient.

As shown in Table 4, the four census variables account for 75 per-cent of the variance in the sixth grade achievement scores.[2] Since this is five points over our goal of 70 percent, we conclude that it is feasible to use census data exclusive of variables referring explicitly to race and ethnicity to account for a substantial proportion of the achievement differences (in grade-level averages) among elementary schools in a big city.

Other related conclusions that emerged from the regression analyses we carried out using various combinations of census and school variables were as follows:

1. Equations using additional census variables referring explicitly to race generally did not account for substantially

[2]Using logs rather than raw scores for the independent (census) variables increased the R^2 to 79 percent. However, logged equations with other dependent (achievement) variables generally resulted in an increase of only 1 to 3 percent in R^2.

more of the variance in achievement scores than did equations using only the first four census variables shown in Table 4. For example, when Per Cent Negro (PCNEG - Table 1, #1) was added to the equation shown in Table 4, the increase in R^2 was only .02 ($p < .005$). When nine other census variables including several designating race were part of a non-forced equation also including three of the four variables shown in Table 4, only two added more than .6 of one percent to the explanatory power of the equation, and the overall R^2 was only .76. Thus, with regard to the goal of predicting elementary school achievement scores in Chicago from census data, there is little to be gained by including variables referring explicitly to race.

2. In almost all non-forced equations regardless of grade level (4th or 6th), type of achievement score (Q2 or Q3), or number of census variables included (in non-forced order), the variables shown in Table 4 either were the first four or were among the first five to enter. Thus it seems clear that these four variables have high utility for predicting grade-level achievement scores among Chicago elementary schools. When the seven best remaining census variables (as indicated by inspection of the correlation matrix) were then entered into a series of non-forced equations utilizing various of the dependent achievement variables, no variable consistently added (in the predicted direction) as much as three percent to the explainable variance. Examples from several of these equations are shown in Appendix A.

3. At the 6th grade level, the census variables shown in Table 4 predicted Q3 scores better than Q2 scores, but the opposite was true at the 4th grade level. Thus there was no reason to conclude that either Q3 scores or Q2 scores are inherently superior to the other in using census data to account for the variance in school achievement means.

4. The percentage of variance in learning ability (I.Q.) scores accounted for by census variables generally was as high or even a little higher than was true with respect to reading achievement scores. For example, the four variables shown in Table 4 accounted for 78 percent of the variance in SLA Q3 scores at the sixth grade level and 71 percent of the variance in SLA median percentile scores at the 4th-grade level. This finding is what one would expect given the premise that school characteristics are likely to have a somewhat greater impact on achievement scores than on more general tests of learning ability.

School Characteristics Variables

The high degree of intercorrelation which exists among census and school variables makes it difficult to draw firm conclusions

about their respective importance and contributions. However, we do have data on Per Cent Entering First Year Ready (EFYRRD - Table 3, #20) and can assume that scores on this variable largely reflect family and neighborhood background rather than school influences since students take the readiness test during their first few months in school.[3] This assumption receives support from a regression analysis showing that the four census variables in Table 4 account for 71 percent of the variance in reading readiness scores among the schools in our sample.

Taking this finding as evidence that the four census variables reflect social background influences associated with readiness to read in schools as presently organized in Chicago, and assuming further that these variables continue to reflect advantages and/or handicaps students experience as they proceed through elementary school, one can conclude that the maximum impact that Chicago elementary schools as currently organized exert in overcoming initial learning disadvantages related to reading is about 25 percent (1 - the 75 percent in sixth grade achievement scores accounted for by the four census variables.)[4] Accepting these assumptions, we can proceed to examine the potency of our school characteristics data in accounting for achievement differences across schools by entering the best combination of school characteristics variables in a regression equation following the four census variables shown in Table 4. The results of this forced-order equation are shown in Table 5.[5]

TABLE 5. Regression Coefficients for Four Selected School Variables Entered in a Sixth Grade Q3 Regression Equation Following the Four Selected Census Variables Shown in Table 4*

Variable	Regression Coefficient	Standard Error	F_2	
FAC6PL	.315	.057	30.878	(<.001)
TAPP	-.217	.337	.414	(<.60)
PCCAUT	.084	.053	2.487	(<.15)
PCCAUC	.059	.047	1.54	(<.30)

*See footnotes Table 4.

[3] It is true that many youngsters attend pre-schools or nurseries and most attend kindergarten, but with the advent of Headstart and other inner city pre-school programs, there is little reason to believe that readiness scores reflect differential access to pre-school resources among differing social groups in a big city.

[4] On the one hand it should be kept in mind that this conclusion does not negate the possibility that some of the schools may be much more effective than others in overcoming the learning handicaps associated with disadvantaged social status. On the other

The data shown in Table 5 indicate that Per Cent Faculty with Six Years or More Experience is the best school variable (in our data set) in accounting for differences in achievement scores among the schools in our sample. When account is taken simultaneously of the four census variables and the four school variables shown in the table, only FAC6PL has a regression coefficient significantly related to Q3 achievement scores in the sixth grade. This conclusion is also supported by the fact that Per Cent Faculty with Six Years or More Experience had higher zero-order correlations with achievement scores than any of the other school characteristics variables. For example, the zero-order correlation between this variable and sixth-grade Q3 achievement was .81; the next highest correlation between a school variable (other than alternate measures of faculty experience) and the Q3 achievement in the sixth grade was .60.

It should be emphasized that our data do not necessarily indicate that faculty experience has more influence on school achievement than do per pupil expenditure, class size, school size, or the other school variables for which we had accessible measures. For one thing, data on the standard deviations of the school characteristics (Table 3) suggest that the Chicago schools have been more successful in equalizing the allocation of other resources across schools than is true with respect to faculty characteristics; thus the other school variables predictably are less closely associated with achievement than the faculty variables. Actually, the school variables are so highly intercorrelated and share so much common variance that we cannot infer causal relations between any given school characteristic and achievement differences among the schools in our sample, particularly since several studies addressed specifically to this question have found no significant relationship between faculty experience and student achievement.[6]

Achievement and Social Background in Black Schools.

Since fifty schools in our sample were 98 to 100 percent black, it was possible to inquire whether neighborhood (census) variables function in the same manner and with equal po ency in predicting school achievement scores in black neighborhoods as

---

hand it should be noted that the 25 percent figure probably is an overestimate since there undoubtedly is some error in our measures of home and neighborhood influence on achievement.

[5] By "best combination of variables, we mean variables which visual inspection of the correlation matrix indicated were highly correlated with a dependent variable but relatively little intercorrelated among themselves.

[6] e.g., see James Fey, "Classroom Teaching of Mathematics," Review of Educational Research, v. 39, no. 4 (October 1969), 535-551; James S. Coleman, et al., Equality of Educational Opportunity. (Washington, D. C.: U. S. Government Printing Office, 1966).

in the sample as a whole. Accordingly, we separated these fifty schools into a separate sub-sample and repeated many of the regression equations described in the preceding sections.[7]

Table 6 shows data obtained using the same forced-order independent and dependent variables as were reported for the total sample in Table 4.

TABLE 6. Per Cent of Variance in Sixth Grade Q3 Scores in Black Schools Accounted for by Four Selected Census Variables*

Variable	R	R^2	Increase in R^2	F_1		Regression Coefficient	Standard Error	F_2	
PCFSEP	.649	.421	-	-		-.016	.681	.001	(**)
PCLACK	.652	.424	.004	.33	(<.60)	-.933	.573	2.654	(<.15)
PC6PER	.738	.545	.120	12.1	(<.005)	-.714	.179	15.918	(<.001)
PCOWN	.759	.576	.031	3.3	(<.10)	.214	.119	3.246	(<.10)

*See footnotes, Table 4. **Purely chance relationship.

These data show that the four census variables account for 58 percent of the variance in Q3 school achievement scores among sixth graders at the fifty black schools in the sub-sample, as compared with 75 percent for the 122-school sample as a whole. More precisely, the data further show that the three variables Per Cent of Families Which Lack One or More Plumbing Facilities; Per Cent of Units with 6 Persons or More, and Per Cent Owner Occupied of Total Occupied Units account for all of the variance explained (58%) in achievement; Per Cent of Females Separated is not associated with achievement level in our sub-sample of black schools when included in the equation with the other three census variables, for which the regression weights are very similar to the corresponding weights in the sample as a whole (Table 4).

In general, this same pattern was repeated at other grade levels and using other census variables and achievement measures. That is, the percentage of variance accounted for by a group of census variables generally was fifteen to twenty-five percentage points lower in the black sub-sample than in the sample as a whole.

[7] Due to the smaller size of the sample, however, we decided not to use more than four or five independent variables in a single equation.

This pattern raises the possibility that census variables may not be as powerful predicting achievement levels in black schools as in predominantly-white schools in Chicago. Unfortunately, there were not enough schools in our sample with 90 per cent or more white students to test this possibility directly with a sub-sample of white schools.

Lacking a comparable sub-sample of white schools, we cannot accept the conclusion that census variables (and the underlying social forces they represent) are less important predictors of achievement in black neighborhoods in Chicago than in white schools or in schools throughout the city as a whole. For one thing, the census variables shown in Table 6 account for nearly half (48%) of the variation in first grade readiness scores among the fifty black schools in our sub-sample, even though the best census predictor of achievement in the total sample (PCFSEP) is unrelated to this dependent variable for the sub-sample (as also is true with respect to sixth-grade Q3 scores). More important, the facts that there are less than half as many cases available for regression analysis and that the range of predictor scores is much less for the sub-sample than the sample as a whole reduce the possibilities of accounting for as high a percentage of variation in the sub-sample as in the total sample. This difference between the black sub-sample and the sample as a whole may well account for most of the fifteen-to-twenty-five point difference we generally found in predicting school achievement scores from selected census variables.

As noted above, the data in Table 6 show that the regression coefficients for the census variables using the sample of black schools are similar to the corresponding coefficients using the 122-school total sample (Table 4), except that the regression weight for Per Cent of Females Separated (PCFSEP) is now insignificant in the full equation. The insignificant regression weight for PCFSEP indicates that this variable is no longer independently related to achievement in the black schools when part of a regression equation including the other three census variables, even though it is the census variable most highly correlated with Q3 sixth-grade achievement ($r = -.79$) in our total sample and is correlated with achievement at -.65 in our sub-sample of black schools.

This finding means that the common variance of the predictor set is somewhat more "unidimensional" in black schools than in the entire sample. The substantive meaning of this finding is somewhat speculative, but empirically it indicates that the indices of poverty and social disorganization are comparatively redundant in the black sub-sample, each variable reflecting more nearly the same set of conditions as reflected by other members of the set. In order to determine whether or not there is a "compounding (i.e., multiplicative) effect asscciated with

the phenomena reflected in common in our measures of poverty and social disorganization in the black sub-sample, in the future we intend to test several equations including product terms using PCFSEP multiplied separately against the other three census variables. The results of such equations of course will embody primarily empirical speculation because of the small number of schools in the black sample, which means that one or two "outlying" schools might substantially affect the results. However, the results of this type of analysis might well prove very valuable in suggesting directions for future research.

To explore this pattern further, we examined additional regression equations using other achievement measures. In each case, whether using Q2 or Q3 scores as the dependent variable, it was found that the regression coefficients for PCFSEP had insignificant F values and accounted for none of the variance in achievement when account also was taken of the other three census variables.

Examination of these equations also showed, however, that until the variable Per Cent Owner Occupied ot Total Occupied Units (PCOWN) was entered in the series of equations, PCFSEP generally had a high negative coefficient. For example, for Q3 sixth-grade achievement the coefficient for PCFSEP followed only by PC6PER and PCLACK was -1.107 ($p < .005$). Comparable coefficients for PCFSEP using Q2 scores in the fourth grade and reading readiness scores as dependent variables were -.98 ($p < .01$) and -1.20 ($p < .025$), respectively. This finding reflects the high redundancy ($r = -.87$) between PCFSEP and PCOWN in our sub-sample of black schools. In black neighborhoods with a high percentage of owner-occupied units, there are relatively low percentages of separated females and achievement is relatively high. In black neighborhoods with a high percentage of renters, there are high percentages of separated females and achievement is low. Consequently, the percentage of separated females does not add as much in predicting achievement as is true in the sample as a whole, for which the zero-order correlation between PCFSEP and PCOWN is only .66.

The fact that PCFSEP has a significant coefficient in equations with the four census variables for the total sample (e.g., Table 4) but not for the black sub-sample indicates that for the sample as a whole its relationship with achievement probably reflects its association with race and, possibly, other variables correlated with poverty and socioeconomically disadvantaged status in Chicago.

The high degree of intercorrelation which exists between renter-occupied units and separated females in our sub-sample of black schools leads us to suspect that the relatively low achievement level in black neighborhoods with high percentages of renters may be reflecting patterns of social disorganization which in turn may be characteristic of low-status neighborhoods with many

families headed by females and little home ownership. The data do not allow us to test this proposition or to determine whether the high incidence of separated females in these neighborhoods may be contributing to or reflecting social disorganization, or some combination of the two. Nor are we able to say whether families with separated females in ghetto neighborhoods with high percentages of renter occupancy may produce children lower in achievement than intact families in these neighborhoods or whether heavily-rental ghetto neighborhoods with many separated females produce relatively low-achieving children regardless of the type of family.

Our data definitely do indicate, however, that density of population is associated with low achievement in predominantly-black schools in Chicago. As shown in Table 8, the most reliable census variable we identified for purposes of predicting school achievement scores in the black sub-sample was Per Cent of Units with 6 Persons or More; the same pattern was found when achievement scores at other grade levels were used for the dependent variable. This finding supports the conclusion that ghetto neighborhoods with high density of population as measured by persons per dwelling unit have lower-achieving children than ghetto neighborhoods with fewer persons per dwelling unit. Whether density has an effect over and beyond its association with socioeconomic level cannot be determined here since we have no usable direct measures of income and occupation from the Third Count data.

Census Variables Not Associated with Race

As mentioned above, many census variables such as Per Cent of Females Separated which are useful in accounting for variation in school achievement levels also are closely associated with race and thus probably reflect aspects of social and family status other than those described in a variable's title. It is very difficult to find social status measures uncorrelated with race in Chicago and most other big cities in the U. S., but one would have to do so if the object were to account for differences in school achievement levels using only student background variables which have no relationship to race.

Since it may be of some interest to determine how much of the variation in achievement scores among the schools in our sample can be accounted for using variables that are mostly independent of race, we identified those which were not closely associated with race but were highly correlated with achievement and computed their multiple correlation with achievement. As might be expected based on preceding analyses, the two best variables for this purpose turned out to be Per Cent of Families

Which Lack One or More Plumbing Facilities (PCLACK) and Per Cent Owner Occupied of Total Occupied Units (PCOWN). None of our other census variables were sufficiently little correlated with race and/or PCLACK and PCOWN and also sufficiently correlated with achievement to allow for entry into more powerful regression equations without also substantially increasing the contamination between independent variables and race.

The multiple correlation using PCOWN and PCLACK as independent variables and Per Cent Negro as the dependent variable is only .33 (R^2 = .11), while the corresponding multiple correlation using Q3 sixth grade scores as the dependent variable is .69 (R^2 = .47). Corresponding multiple correlations using first and fourth grade data also were computed and were found to be .75 (R^2 = .56) using readiness scores as the dependent variable and .69 (R^2 = .48) using Q2 fourth grade scores as the dependent variable. These results indicate that PCOWN and PCLACK have considerable power independent of race in accounting for the variation in achievement levels among Chicago elementary schools and also that the selection of the two variables for this purpose was not merely the result of a fortuitous correlation matrix at only one grade level.

Concentrated Urban Poverty

It will be recalled that one of the major hypotheses to be investigated in this study was that low-status children in densely-populated urban poverty areas achieve more poorly in the school than do similar-status children in neighborhoods thought to be less characterized by social disorganization. One test we proposed for investigating this hypothesis was to determine whether interaction, squared, and cubed regression terms would add appreciably to the prediction of achievement after account was taken of social background, as had proven to be the case in our reanalysis of the data published by Garms and Smith (see chapter 1).

We did not find cases in which the cross-product, squared, and/or cubed terms (see Table 2) contributed substantial variance in "explaining" achievement scores in predicted directions after account was taken of some or all of the four census variables shown in Table 4. However, we do not feel that this finding invalidates the hypothesis, because two of the four variables in question are themselves indicators of social disorganization (i.e., PCFSEP) and density of population (i.e., PC6PER) in a big city and the other two (i.e., PCLACK and PCOWN) probably are associated with poverty. From this point of view our data do suggest that density of population and social disorganization are associated with low school achievement over and beyond the

association between achievement and poverty per se.[8] Data involving the interaction, squared, and cubed terms did indicate, however, that the relationships between these variables and achievement may be more complicated than we had expected based on reanalysis of the New York data.

Exploring these relationships, we found that at least two of our census variables seemed to make a contribution in accounting for achievement variance in a direction opposite to that we had predicted. These two variables were Per Cent of Occupied Units with Female-Head Families (PCFHED - Table 1, #14) and Per Cent of Occupied Housing Units with 1.51 or More Persons Per Room x Per Cent of Occupied Units with Female Headed Families (PC151 x PCFHED - Table 2, #12). Even when account was taken through regression analysis of the four census variables shown in Table 4, in separate equations these two variables contributed from two-to-six per cent more variance and had positive regression coefficients which were significant at or beyond the .01 level regardless of the grade level (1st, 4th, or 6th) or type of achievement score (Q2 or Q3).

This finding may be entirely spurious in that it may be caused either by multicollinearity[9] or by a few outlying schools which for some reason have unexpectedly high or low achievement or are mis-classified in terms of the census data. On the other hand it also may mean that other background characteristics being equal, schools in which there is a high percentage of female-headed families and/or a high density of persons per room in such families may have slightly higher achievement than schools with opposite characteristics.[10] Since preliminary analysis

---

[8] Unfortunately we have no direct measure of poverty to test this hypothesis more directly. We hope to explore this issue further in subsequent analyses with Third and Fourth Count data.

[9] The term "multicollinearity" refers to a situation in which there is high intercorrelation among independent variables.. As has been shown by Gordon, such variables have unstable regression coefficients and sometimes produce spurious reversals of sign in a regression analysis which are not reproducible even in a second sample from the same population. (See Robert A. Gordon, "Issues in Multiple Regression," The American Journal of Sociology, v. 73 /1968/, 596) However the fact that we find these reversals at three different grade levels and in a variety of regression equations using differing census variables leaves open the possibility that the finding may not be entirely spurious.

[10] Such a pattern theoretically could occur if public housing projects in poverty areas had higher percentages of female-headed families and also slightly less social disorganization or slightly more effective schools than is true in the surrounding blocks.

indicated that similar trends existed among the sub-sample of black schools (where the percentage of female-headed families was greater than in the total sample), we decided to explore the issue further by adding these two variables to regression equations including the three variables (PCOWN, PC6PER, and PCLACK) previously identified as being independently associated with achievement in this sub-sample (Table 6).

The results of these regression analyses showed that Per Cent of Occupied Housing Units with 1.51 or More Persons Per Room x Percent of Occupied Units With Female Headed Families was neither significantly related to achievement nor added to the prediction of achievement levels with the other variables controlled, and Per Cent of Occupied Units with Female Head Families (PCFHED) added only a few percentage points in accounting for variance in achievement levels in the sub-sample of black schools. For example, PCFHED added only two per cent in accounting for the variance in Q3 achievement scores at the sixth grade level when entered into a regression equation following PCOWN, PC6PER, and PCLACK. Because this result could have been due to the influence of inaccurate data on one or two schools and/or to the spurious effects of multicollinearity, we conclude that the analysis does not provide any real support for the possibility that percentage of female-headed families may be positively associated with achievement when account is taken of other family and neighborhood variables highly correlated with achievement levels in the Chicago Public Schools.

This conclusion was further supported by a plot of the achievement residuals against PCFHED. This plot showed that the three schools which had the highest percentage of separated females in their neighborhoods[11] also scored from two to twelve points less than predicted on achievement. This result makes it unlikely that the true relationship between PCFHED and achievement is positive in our sample of fifty black schools.

Student Body Race and Ethnicity in Predicting Achievement

As mentioned above, the data do not allow us to determine with any degree of certainty whether or not achievement scores among the schools in our sample are improved by racial and/or socio-economic integration of their student bodies. Though we cannot use our data to answer this important question, it may be of some interest to consider whether race and ethnicity contribute to the prediction of achievement when account is taken of the census variables we have identified as being closely related to readiness scores in the first grade and achievement in the later grades.

[11] The three schools had percentages approximately fifty per cent higher than the next highest school.

To do this, Table 7 shows data on the relationship between Q3 achievement at the sixth grade and variables denoting four of the major racial and ethnic groups in the Chicago public schools: Caucasian (non-Spanish-speaking); African; Mexican; and Puerto Rican. The table shows the increased percentage of variance accounted for and data on the regression coefficients when each of the four racial-census variables is added separately to the regression equation including the four census variables shown in Table 4.

TABLE 7. Contribution to Variance Explained in Q3 Sixth Grade Scores and Regression Coefficients of Four Racial-Ethnic Variables Added Separately to Equation Including Four Best Census Variables*

Variable Added	Increase in R^2	F_1	Regression Coefficient	Standard Error	F_2
PCCAUC	.050	28.9 ($<.001$)	.178	.033	28.922 ($<.001$)
PCAFRI	.029	15.2 ($<.001$)	-.143	.037	14.907 ($<.001$)
PCMEX	.004	1.9 ($<.25$)	-.108	.078	1.9503 ($<.15$)
PCPUER	.006	2.8 ($<.10$)	-.153	.094	2.669 ($<.15$)

*The four census variables are shown in Table 4. See Table 4 footnotes.

The pattern in Table 7 indicates that per cent white (excluding Spanish-speaking background) is positively and significantly related to achievement, per cent black is negatively and significantly related to achievement, and per cent Mexican and Puerto Rican may be slightly related to achievement in a negative direction, when account is taken of socioeconomic background as measured by the four census variables.

When similar tables were constructed using fourth and sixth grade Q2 scores and fourth grade Q3 scores, essentially the same pattern emerged except that per cent Puerto Rican was significantly related ($p < .05$) to achievement using Q2 scores and per cent black was not significantly related to achievement using Q2 sixth grade scores.

In interpreting these data, it should be kept in mind that the range of scores on per cent Mexican and per cent Puerto Rican was considerably less than on per cent white and per cent black. That is, many more schools had no Mexican and/or Puerto

Rican students than had no black and/or white students. It is possible that the former two variables would have been more clearly related to achievement if Mexican and Puerto Rican students had been represented among more of the schools in the sample.

Even so, however, the data suggest that adding variables showing the respective percentages of minority groups (i.e., black, Mexican, Puerto Rican) to an equation including the four census variables from Table 4 does not contribute a great deal in accounting for variation in school achievement scores in the Chicago Public Schools: the percentage of additional variance accounted for ranges from a low of almost zero to a high of about five of six points, and even the latter figure certainly reflects to some extent the influence of other intercorrelated variables denoting school characteristics and additional census variables. In view of the political problems involved in allocating compensatory education funds on the basis of race or ethnicity, this finding in turn suggests that it would be about as satisfactory to take account only of socioeconomic background variables as to include variables directly denoting minority racial or ethnic status. Just as it would be useful for educators to help the public understand that socioeconomic segregation rather than racial segregation per se appears to be the more important factor in depressing achievement in the public schools,[12] emphasizing socioeconomic considerations in the allocation of compensatory education funds might reduce the tendency for people to confuse race or ethnicity with social class in considering ways to improve achievement levels in urban schools.

Feasibility of Using Census Data for Allocating Compensatory Education Funds

The main object of this study was to determine whether census data not directly denoting race or ethnicity in school neighborhoods could be used relatively easily to account for 70 per cent or more of the grade-level achievement scores among the public schools of a big city. This goal was achieved when it was found that only four variables (Percent of Females Separated; Per Cent of Families Which Lack One or More Plumbing Facilities; Per Cent Owner Occupied of Total Occupied Housing Units; and Per Cent of Units with 6 Persons or More) accounted for 65 to 75% of the variance in grade-level averages no matter whether fourth- or sixth-grade scores and Q2 or Q3 scores were used as the dependent variable.

---

[12] See James S. Coleman et al., op. cit., and Alan B. Wilson, The Consequences of Segregation (Berkeley: Glendessary, 1969).

One reason why there is widespread interest in this type of study is because of the possibility that census data might be used to allocate funds to improve educational opportunities for students whose socioeconomic background makes them educationally disadvantaged in most existing public school systems. At present, for example, federal funds for compensatory education under Title I of the Elementary and Secondary Education Act are allocated according to and must be expended on students whose families have incomes below officially-designated poverty levels. Many educators consider this procedure undesirable and even counterproductive because it singles out individual students who tend to feel stigmatized by the poverty designation and also because it requires a great deal of possibly inaccurate and burdensome paperwork to collect income and related information on every student. Easily-collected census information might serve equally or better in allocating funds for compensatory education than do the cumbersome and questionable procedures now in use.

Before concluding, a few cautions also are in order for other researchers who may wish to replicate our findings elsewhere:

1. We made a definite attempt to omit schools from our sample if we had strong reasons to believe that a school differed substantially from its neighborhood. For example, we did not use schools along the lake shore where there are expensive, high-rise apartment developments with families that typically have no school-age children. This phenomena probably should be taken into account in other cities.

2. We did not include elementary schools in our sample unless they had all the grades from one through six. It is possible that inclusion of "lower-grade" centers or other organizational variants might have modified our results somewhat.

As usual in an exploratory study of this type, there are several additional questions which would have to be answered by additional research before an adequately-informed decision could be made about the point at issue. Among these are the following:

1. Does the "best" set of census variables (Table 4) we identified in the present study--or some equally-accessible set--function with the same potency and stability in predicting achievement scores in other cities and other types of communities?

2. To what degree do 1970 census data lose utility in predicting achievement scores collected three or five or eight years in the future as compared with the 1970-1971 achievement scores used in the present study?

3. Do variables from the Fourth Count of the Census, which includes income and occupational data but by tracts rather than blocks, have equal or greater utility in accounting for grade-level achievement scores as do the Third Count variables used in the present study?

4. Do 1970 census data allow for as good or better classification of schools as do methods currently in use for determining eligibility for federal funding under Title I of the Elementary and Secondary Education Act? That is, is the multiple correlation between the best set of census variables and achievement as high or higher as the current correlation between achievement and school-level participation in Title I?

5. What accounts for the appearance of a few highly discrepant schools in our sample? (A few "discrepant" schools have achievement much higher or lower than predicted by the four-census-variable equation.) Do these schools have instructional programs which for some reason are more or less effective than other schools? Do their staff members take particular pains to "teach" toward the tests? Alternately, is their student body composition for some reason much different from the neighborhood composition? If the latter explanation is correct, how much would their elimination improve the explanatory power of the regression equation?

6. Can the census data be used to identify different types of neighborhoods (i.e., in addition to the differences already considered in this paper)? If so, can predictions of achievement be significantly improved by taking type of neighborhood into account?

We hope to explore these questions in subsequent research utilizing 1970 census data to determine how useful easily-accessible information on urban neighborhoods may be for "explaining" differences in achievement levels among public schools and for related purposes such as allocating special funds to compensate for the disadvantaged background of many students in low-achieving schools.

APPENDIX A

Per Cent of Additional Variance in Sixth Grade Q2 and Q3 Scores Accounted for by Selected Census Variables Entered Separately in Forced-Order Equations Following the Four Best Census Variables Identified in TABLE 4.

Variable	Percent of Additional Variance Using Q2 Scores	Percent of Additional Variance Using Q3 Scores
PCLACK (Per cent of Population in Units Lacking 1 or More Plumbing Facilities and with 1.01 or More Persons Per Room)	.001	.000
PCLACK Squared	.010	.005
PCFHED (Per cent of Occupied Units with Female Head Families)	.020*	.029*
PC151 (Per cent of Occupied Housing Units 1.51 or More Persons Per Room)	.004	.002
PCNEG (Per cent Negro)	.041	.022
PC151 x PCFHED	.043*	.044*

* Positive regression coefficient